

# Constrained Subspace Estimation via Convex Optimization

I. Santamaria, J. Via  
Dept. Ingeniería de Comunicaciones  
Universidad de Cantabria  
Plaza de la Ciencia, Santander, 39005, Spain

M. Kirby, T. Marrinan, C. Peterson, L. Scharf  
Department of Mathematics  
Colorado State University  
Fort Collins, CO, 80523, USA

**Abstract**—Given a collection of  $M$  experimentally measured subspaces, and a model-based subspace, this paper addresses the problem of finding a subspace that approximates the collection, under the constraint that it intersects the model-based subspace in a predetermined number of dimensions. This constrained subspace estimation (CSE) problem arises in applications such as beamforming, where the model-based subspace encodes prior information about the direction-of-arrival of some sources impinging on the array. In this paper, we formulate the constrained subspace estimation (CSE) problem, and present an approximation based on a semidefinite relaxation (SDR) of this non-convex problem. The performance of the proposed CSE algorithm is demonstrated via numerical simulation, and its application to beamforming is also discussed.

**Index Terms**—Subspace averaging, Grassmann manifold, convex optimization, semidefinite relaxation.

## I. INTRODUCTION

Input data in many signal processing applications admit a subspace representation, i.e., the intrinsic dimension of the objects of interest is much smaller than the dimension of the ambient space in which they live. Examples include image and video recognition where a single object under different pose or illumination conditions can be modeled as a subspace [1], or pattern recognition applications in which features obtained after a dimensionality reduction stage such as principal component analysis (PCA) are commonly used [2]. In other signal analysis problems, one has to cope with the effects of the multiplicity of appearances of objects under different warpings or elastic deformations, where an invariant representation of the signal is shown to have the form of a subspace [3].

A central problem in these applications is to estimate an average subspace from a collection of experimentally measured subspaces. In our previous work [4], [5], we have presented a solution for this subspace averaging problem that minimizes an extrinsic mean-squared error, defined by the squared Frobenius norm between projection matrices. In addition, the solution in [4] automatically returns the dimension of the optimal average subspace. Variations on this problem have been considered in [6] and [7]. An application to image classification that first estimates the mean representation of the object manifold by means of the proposed subspace averaging technique has recently been described in [8].

In this paper, we extend our previous work and consider the problem of finding a subspace that best approximates the collection of subspaces according to an extrinsic distance measure and, at the same time, intersects a model-based subspace in a given number of dimensions. The model-based subspace can be determined from the physical process that generates the measurements, or from prior information about the problem. For instance, in a beamforming problem [9] the model-based subspace might encode prior information about the angle of arrival of a source impinging on a linear array of antennas. When the direction-of-arrival (DoA) of the impinging source is not perfectly known because of mismatch errors due to imperfect array calibration or local scattering problems, this prior information can be encoded in the form of a low-dimensional subspace in which the incident signal is expected to lie.

The rest of the paper is organized as follows: in Section II, we provide a brief review of the subspace averaging technique. Section III formulates the non-convex constrained subspace estimation (CSE) problem, and presents an approximate solution based on semidefinite relaxation (SDR). In Section IV we illustrate the performance of the CSE algorithm by means of numerical simulations, and present an application to beamforming as a motivating example. The main conclusions are summarized in Section V.

**Notation** The notation  $\mathbb{G}(d, n)$  will be used to represent the complex Grassmann manifold whose points parameterize all subspaces of dimension  $d$  in an  $n$ -dimensional ambient space. We use  $\langle \mathbf{A} \rangle \in \mathbb{G}(d, n)$  to denote a point (subspace) on the Grassmann manifold, whereas  $\mathbf{A}$  is used to denote a matrix whose columns form a unitary basis for that subspace. The superscripts  $(\cdot)^T$  and  $(\cdot)^H$  denote transpose and Hermitian, respectively. The trace, rank, and Frobenius norm of a matrix  $\mathbf{B}$  will be denoted, respectively, as  $\text{tr}(\mathbf{B})$ ,  $\text{rank}(\mathbf{B})$ , and  $\|\mathbf{B}\|_F$ . Finally,  $\mathbf{I}_n$  denotes the  $n \times n$  identity matrix.

## II. SUBSPACE AVERAGING

Given a collection of experimentally measured subspaces  $\langle \mathbf{V}_m \rangle \in \mathbb{G}(q_m, n)$ ,  $m = 1, \dots, M$ ; a classical problem in statistical signal processing is that of finding an average subspace of dimension  $d_s$ ,  $\langle \mathbf{V}_s \rangle \in \mathbb{G}(d_s, n)$ , that “best approximates” the collection of subspaces according to an extrinsic distance measure.

The extrinsic distance metric between two subspaces,  $\langle \mathbf{V}_s \rangle$  and  $\langle \mathbf{V}_m \rangle$ , is defined as the Frobenius norm of the difference between the respective projection matrices [10], [5]

$$d(\langle \mathbf{V}_s \rangle, \langle \mathbf{V}_m \rangle) = \|\mathbf{P}_s - \mathbf{P}_m\|_F, \quad (1)$$

where  $\mathbf{P}_s = \mathbf{V}_s \mathbf{V}_s^H$  and  $\mathbf{P}_m = \mathbf{V}_m \mathbf{V}_m^H$ . Notice that the two subspaces can have different dimensions. Actually, writing the squared extrinsic distance in terms of the cosines of the principal angles between the two subspaces it is easy to show that [11]

$$\|\mathbf{P}_s - \mathbf{P}_m\|_F^2 = q_m + q_s - 2 \sum_{r=1}^{\min(q_m, q_s)} \cos^2(\theta_r). \quad (2)$$

In our previous work [4], [5], we showed that, given the collection  $\{\langle \mathbf{V}_m \rangle\}_{m=1}^M$ , the average or central subspace of dimension  $d_s$  maximizes the cost function

$$\max_{\mathbf{V}_s \in \mathbb{S}(n, d_s)} \text{tr}[\mathbf{V}_s^H \bar{\mathbf{P}} \mathbf{V}_s], \quad (3)$$

where  $\mathbb{S}(n, d_s)$  denotes the complex Stiefel manifold of orthonormal  $d_s$ -frames in  $\mathbb{C}^n$ , and  $\bar{\mathbf{P}}$  is the average of projection matrices, defined as

$$\bar{\mathbf{P}} = \frac{1}{M} \sum_{m=1}^M \mathbf{P}_m. \quad (4)$$

This average is not, itself, a projection matrix. It is known that the solution of (3) is given by any unitary matrix whose column space is the same as the subspace spanned by the  $d_s$  principal eigenvectors of  $\bar{\mathbf{P}}$  [12], [13]. Moreover, in [4] it was shown that the optimal choice of  $d_s$  is determined by the index of the last eigenvalue of  $\bar{\mathbf{P}}$  that exceeds 1/2, a result that is identical to a finding in [14] for the unrelated problem of constructing low-dimensional time-frequency projections.

### III. CONSTRAINED SUBSPACE ESTIMATION

In some applications, in addition to the collection of measured subspaces we might have a model-based subspace  $\langle \mathbf{W} \rangle$  of dimension  $d_w$  that is used to constrain the sought solution by requiring the solution to intersect  $\langle \mathbf{W} \rangle$  in  $d_i$  dimensions. This model-based subspace can be determined from the physical process that generates the measurements, or from prior information about the problem. As a motivating example, consider for instance a beamforming problem in which one of the impinging signals is known to lie in a given subspace of dimension  $d_w$  (as we will see in Section IV this model-based subspace can be determined from the DoA of the estimated source and the presumed mismatch error).

**Problem Statement.** Our problem can be stated as follows: given a collection of experimentally measured subspaces  $\langle \mathbf{V}_m \rangle \in \mathbb{G}(q_m, n)$ , for  $m = 1, \dots, M$ , and a model-based subspace  $\langle \mathbf{W} \rangle \in \mathbb{G}(d_w, n)$ , find an average subspace  $\langle \mathbf{V}_s \rangle$  of dimension  $d_s$  which maximizes the cost function of (3), under the constraint that it intersects the model-based subspace in  $d_i$  dimensions. To avoid trivial cases where the unconstrained solution is optimal, we assume  $d_s + d_w \leq n + d_i$ .

From the unconstrained subspace average described in Section II, it is clear that the information from the collection of subspaces can be summarized through the average projection matrix  $\bar{\mathbf{P}}$ . Therefore, the optimization problem we want to solve can be written as

$$\begin{aligned} & \text{maximize}_{\mathbf{V}_s \in \mathbb{C}^{n \times d_s}} \text{tr}(\mathbf{V}_s^H \bar{\mathbf{P}} \mathbf{V}_s) \\ & \text{subject to} \quad \mathbf{V}_s^H \mathbf{V}_s = \mathbf{I}_{d_s}, \\ & \quad \dim(\langle \mathbf{V}_s \rangle \cap \langle \mathbf{W} \rangle) \geq d_i. \end{aligned} \quad (5)$$

The solution can be parameterized as follows

$$\mathbf{V}_s = \left[ \begin{array}{c|c} \mathbf{W} & \mathbf{W}^\perp \\ \hline \mathbf{0} & \mathbf{B} \end{array} \right], \quad (6)$$

where  $\mathbf{W} \in \mathbb{C}^{n \times d_w}$  is a matrix whose columns form a unitary basis for the model-based subspace and  $\mathbf{W}^\perp \in \mathbb{C}^{n \times (n-d_w)}$  is a unitary basis for its orthogonal complement. The matrices  $\mathbf{A} \in \mathbb{C}^{d_w \times d_i}$  and  $\mathbf{B} \in \mathbb{C}^{n \times (d_s - d_i)}$  are unitary slices to be determined during the optimization procedure. It is clear that any matrix that can be expressed as (6) determines a subspace of dimension  $d_s$  that intersects  $\langle \mathbf{W} \rangle$  in at least  $d_i$  dimensions.

Let us define

$$\bar{\mathbf{W}} = \left[ \begin{array}{c|c} \mathbf{W} & \mathbf{W}^\perp \end{array} \right],$$

and

$$\bar{\mathbf{X}} = \bar{\mathbf{W}}^H \bar{\mathbf{P}} \bar{\mathbf{W}} = \begin{bmatrix} \bar{\mathbf{X}}_{11} & \bar{\mathbf{X}}_{12} \\ \bar{\mathbf{X}}_{21} & \bar{\mathbf{X}}_{22} \end{bmatrix},$$

where  $\bar{\mathbf{X}}_{11}$  is a  $d_w \times d_w$  matrix. With these definitions, problem (5) can be re-phrased as

$$\begin{aligned} & \text{maximize}_{\mathbf{A}, \mathbf{B}} \text{tr}(\mathbf{A}^H \bar{\mathbf{X}}_{11} \mathbf{A}) + \text{tr}(\mathbf{B}^H \bar{\mathbf{X}} \mathbf{B}) \\ & \text{subject to} \quad \mathbf{A}^H \mathbf{A} = \mathbf{I}_{d_i}, \\ & \quad \mathbf{B}^H \mathbf{B} = \mathbf{I}_{d_s - d_i}, \\ & \quad [\mathbf{A}^H | \mathbf{0}] \mathbf{B} = \mathbf{0}. \end{aligned} \quad (7)$$

Now, taking into account that

$$\text{tr}(\mathbf{A}^H \bar{\mathbf{X}}_{11} \mathbf{A}) = \text{tr}(\bar{\mathbf{X}}_{11} \mathbf{A} \mathbf{A}^H) = \text{tr}(\bar{\mathbf{X}}_{11} \mathbf{P}_A)$$

where  $\mathbf{P}_A$  is a projection matrix defined by

$$\begin{aligned} \mathbf{0} & \leq \mathbf{P}_A \leq \mathbf{I}_{d_w}, \\ \text{tr}(\mathbf{P}_A) & = d_i, \\ \text{rank}(\mathbf{P}_A) & = d_i, \end{aligned}$$

we can rewrite (7) as

$$\begin{aligned} & \text{maximize}_{\mathbf{P}_A, \mathbf{P}_B} \text{tr}(\bar{\mathbf{X}}_{11} \mathbf{P}_A) + \text{tr}(\bar{\mathbf{X}} \mathbf{P}_B) \\ & \text{subject to} \quad \mathbf{0} \leq \mathbf{P}_A \leq \mathbf{I}_{d_w}, \quad \mathbf{0} \leq \mathbf{P}_B \leq \mathbf{I}_{d_n}, \\ & \quad \text{tr}(\mathbf{P}_A) = d_i \quad \text{tr}(\mathbf{P}_B) = d_s - d_i, \\ & \quad \text{rank}(\mathbf{P}_A) = d_i \quad \text{rank}(\mathbf{P}_B) = d_s - d_i, \\ & \quad \left[ \begin{array}{c|c} \mathbf{P}_A & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] + \mathbf{P}_B \leq \mathbf{I}_n, \end{aligned} \quad (8)$$

where the last constraint enforces the orthogonality between the subspaces spanned by the basis  $[\mathbf{A}^T | \mathbf{0}]^T$  and  $\mathbf{B}$ .

Except for the rank constraints, problem (8) is convex. Therefore we can use the idea of semidefinite relaxation (SDR) [15] and solve the following convex relaxation (eliminating also some redundant constraints)

$$\begin{aligned} & \underset{\mathbf{P}_A, \mathbf{P}_B}{\text{maximize}} && \text{tr}(\overline{\mathbf{X}}_{11} \mathbf{P}_A) + \text{tr}(\overline{\mathbf{X}} \mathbf{P}_B) \\ & \text{subject to} && \mathbf{P}_A \succeq \mathbf{0}, \quad \mathbf{P}_B \succeq \mathbf{0}, \\ & && \text{tr}(\mathbf{P}_A) = d_i, \quad \text{tr}(\mathbf{P}_B) = d_s - d_i, \\ & && \begin{bmatrix} \mathbf{P}_A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \mathbf{P}_B \leq \mathbf{I}_n. \end{aligned} \quad (9)$$

To solve problem (9) we used CVX, a package for specifying and solving convex programs [16]. When the relaxation is tight, the solution of (9) returns projection matrices  $\mathbf{P}_A$  and  $\mathbf{P}_B$  of ranks  $d_i$  and  $d_s - d_i$ , and the global optimum satisfies all constraints; otherwise, we need some approximation. There are several ways proposed in the literature to extract solutions with the desired ranks when the problem is not tight, such as Gaussian randomization or rank reduction procedures (also called purification methods) [17]. These methods, however, incur some computational cost. Therefore, in the following subsection we propose an approximate solution to be applied when the problem is not tight.

**Approximate linear algebraic solution to the SDR.** In the convex relaxation, the coefficient matrices  $[\mathbf{A}^T | \mathbf{0}]^T$  and  $\mathbf{B}$  are estimated simultaneously. In fact, problem (9) can be viewed as two max-trace problems which are coupled through the constraints. The approximation proceeds in two steps: first, it obtains the part of the solution that intersects  $\langle \mathbf{W} \rangle$  in  $d_i$  dimensions, and second, it completes the subspace with the remaining  $d_s - d_i$  dimensions which are extracted from the orthogonal complement of the  $d_i$ -dimensional subspace obtained in the first step. In the proposed approximate solution, the two subspaces are obtained sequentially by solving two separate problems.

- 1) Obtain a  $d_i$ -dimensional subspace belonging to the model-based subspace  $\langle \mathbf{W} \rangle$ . To this end solve

$$\begin{aligned} & \underset{\mathbf{A} \in \mathbb{C}^{d_w \times d_i}}{\text{maximize}} && \text{tr}(\mathbf{A}^H \mathbf{W}^H \overline{\mathbf{P}} \mathbf{W} \mathbf{A}) \\ & \text{subject to} && \mathbf{A}^H \mathbf{A} = \mathbf{I}_{d_i} \end{aligned} \quad (10)$$

whose solution,  $\mathbf{A}^*$ , is given by any unitary matrix whose column space is the same as the subspace spanned by the  $d_i$  principal eigenvectors of  $\overline{\mathbf{P}}_{\mathbf{W}} = \mathbf{W}^H \overline{\mathbf{P}} \mathbf{W}$ . Then, the basis for  $\mathbf{W} \mathbf{A}^*$  is

$$\mathbf{F} = \mathbf{W} \mathbf{A}^*$$

- 2) Now, let  $\mathbf{P}_F^\perp = \mathbf{I}_n - \mathbf{F} \mathbf{F}^H$  be the projection onto the orthogonal complement of  $\mathbf{F}$ . Similarly to the previous step, a basis for the remaining portion of the subspace that lives in the space spanned by the columns of  $\mathbf{W}^\perp$  is given by the  $d_s - d_i$  main eigenvectors of the  $(\mathbf{P}_F^\perp)^H \overline{\mathbf{P}} \mathbf{P}_F^\perp$ . We denote this  $n \times (d_s - d_i)$  matrix as  $\mathbf{G}$ .
- 3) A unitary basis for the constrained subspace is finally given by

$$\mathbf{V}_s = [\mathbf{F} \mid \mathbf{G}]. \quad (11)$$

---

**Algorithm 1:** Constrained Subspace Estimation (CSE) algorithm.

---

**Input:** Average of projections  $\overline{\mathbf{P}}$ , unitary basis for the model-based subspace  $\mathbf{W}$ ,  $d_s$ ,  $d_w$ ,  $d_i$

**Output:** Unitary basis for the average subspace  $\mathbf{V}_s$  such that  $\dim(\langle \mathbf{V}_s \rangle \cap \langle \mathbf{W} \rangle) = d_i$

**Initialize:** Generate  $\overline{\mathbf{W}} = [\mathbf{W} \mid \mathbf{W}^\perp]$  and  $\overline{\mathbf{X}} = \overline{\mathbf{W}}^H \overline{\mathbf{P}} \overline{\mathbf{W}}$

**Solve convex problem (9):** Output  $\mathbf{P}_A$  and  $\mathbf{P}_B$

**if Relaxation is tight then**

- 1.- Find an  $d_w \times d_i$  unitary basis for  $\mathbf{P}_A$ :  $\mathbf{F}_1$
- 2.- Find an  $n \times (d_s - d_i)$  unitary basis for  $\mathbf{P}_B$ :  $\mathbf{F}_2$
- 3.- Find solution as

$$\mathbf{V}_s = \left[ \begin{array}{c|c} \mathbf{F}_1 & \mathbf{F}_2 \\ \hline \mathbf{0}_{(n-d_w) \times d_i} & \end{array} \right]$$

**else**

- 1.- Solve problem (10) and find  $\mathbf{F} = \mathbf{W} \mathbf{A}^*$
  - 2.- Find a unitary basis  $\mathbf{G}$  formed by the  $(d_s - d_i)$  main eigenvectors of  $(\mathbf{P}_F^\perp)^H \overline{\mathbf{P}} \mathbf{P}_F^\perp$
  - 3.- Find solution as  $\mathbf{V}_s = [\mathbf{F} \mid \mathbf{G}]$
- 

Finally, the proposed Constrained Subspace Estimation (CSE) algorithm is summarized in Algorithm 1.

#### IV. SIMULATION RESULTS

**Example 1.** We generate  $M$  perturbed versions of a central subspace  $\langle \mathbf{V}_0 \rangle \in \mathbb{G}(d_s, n)$ , as follows:

- 1) Generate a unitary basis for the model-based subspace  $\langle \mathbf{W} \rangle \in \mathbb{G}(d_w, n)$ .
- 2) Generate a unitary basis for the central subspace,  $\langle \mathbf{V}_0 \rangle \in \mathbb{G}(d_s, n)$ , taking at random  $d_i$  dimensions from  $\mathbf{W}$  and  $d_s - d_i$  dimensions from  $\mathbf{W}^\perp$ .
- 3) Generate a collection of  $M$  matrices of dimensions  $n \times n$  as follows:

$$\mathbf{G}_m = [\mathbf{V}_0 | \mathbf{0}] + \sigma \mathbf{Z}_m, \quad m = 1, \dots, M \quad (12)$$

where  $\mathbf{V}_0 \in \mathbb{C}^{n \times d_s}$  is a unitary basis for  $\langle \mathbf{V}_0 \rangle$ , and  $\mathbf{Z}_m \in \mathbb{C}^{n \times n}$  is a matrix whose entries are independent and identically distributed complex Gaussian random variables with zero mean and variance  $1/n$ . The value of  $\sigma$  determines the signal-to-noise-ratio, which is defined as  $\text{SNR} = 10 \log_{10} \left( \frac{d_s}{n \sigma^2} \right)$ . An orthogonal basis for the  $m$ -th subspace,  $\mathbf{V}_m$ , is then constructed from the first  $d_s$  orthonormal vectors of the QR decomposition of  $\mathbf{G}_m$ .

In words, the basis of each measured subspace  $\mathbf{V}_m$  is an additively perturbed version of the basis of the signal subspace  $\mathbf{V}_0$ , which intersects  $\langle \mathbf{W} \rangle \in \mathbb{G}(d_w, n)$  in  $d_i$  dimensions. On the other hand, the SNR measures the spreading of the subspaces: i.e., a low SNR means that the subspaces are very far from each other, whereas a high SNR means that the subspaces are tightly clustered around  $\langle \mathbf{V}_0 \rangle$ .

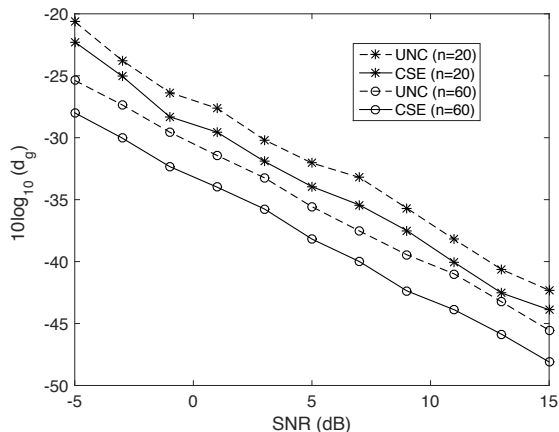


Fig. 1. Extrinsic distance between the true subspace and the estimate provided by the UNC and the CSE algorithms when the ambient space dimension varies ( $n = 20$  and  $n = 60$ ) while other parameters are kept constant ( $d_s = d_w = 8$ ,  $d_i = 4$  and  $M = 10$  subspaces).

For the first example, we have considered a collection of  $M = 10$  subspaces in  $\mathbb{G}(8, 20)$ . The solution and the model-based subspaces both have dimension 8, and they are constrained to intersect in  $d_i = 4$  dimensions. As a figure of merit we compute the extrinsic distance between the solution  $\mathbf{V}_s$  and the true  $\mathbf{V}_0$  defined in (1). For comparison, we also compute the results of the unconstrained average subspace (cf. Section II) which is denoted here as UNC. Fig. 1 shows the average extrinsic distance in logarithmic scale estimated from 100 independent Monte Carlo simulations for the UNC method (in dashed line) and proposed CSE algorithm (in solid line) for two different values of the ambient space dimension. As we can observe, when the subspaces are spread out (i.e., for low SNR values) the CSE algorithm, which enforces the intersection constraint with the model-based subspace in the required number of dimensions, provides a more accurate estimate. Keeping fixed the rest of the parameters, the estimate for both methods improves when the ambient space dimension grows.

**Example 2.** In the second example, we analyze the tightness of the semidefinite relaxation. In general, the tightness depends on a variety of parameters such as the ambient dimension, the number of subspaces, or the SNR. As an example, Fig. 2 shows the tightness estimated from 100 independent Monte Carlo simulations when  $M = 100$ ,  $n = 20$ ,  $d_w = 6$  and  $d_s = 6$ . We can see that for high SNRs (all subspaces are well clustered) the problem is tight with very high probability.

**Application to beamforming.** Let us consider an array of  $n$  sensors and  $K$  uncorrelated sources impinging on the array. Assuming the standard AWG model for the noise, the theoretical covariance matrix of the array output vector has

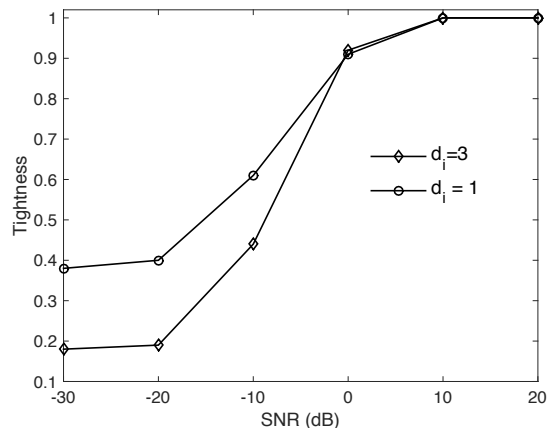


Fig. 2. Tightness for the SD relaxation ( $n = 20$ ,  $d_s = 6$ ,  $d_w = 6$  and  $M = 100$ ). The dimension of the intersection is  $d_i = 3$  or  $d_i = 1$ .

the following form

$$\mathbf{R} = \sum_{k=1}^K \sigma_k^2 \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^H + \sigma^2 \mathbf{I},$$

where  $\sigma_k^2$  are the powers of the incident signals, and  $\mathbf{a}(\theta_k)$  are the steering vectors.

In classical beamforming problems the goal is to estimate the angle or direction-of-arrival (DoA) of every impinging source. In our problem, however, we do not look for accurate DoA estimates, but only for a good approximation (in terms of captured signal power) of the span of the corresponding steering vectors. Suppose now that for a given source (without loss of generality we take this to be the first source) we have a previous estimate of the DoA with a possible angle mismatch error up to  $\Delta$  degrees. Therefore, the true DoA belongs to the interval  $\theta_1 \in \Omega = [\hat{\theta}_1 - \Delta, \hat{\theta}_1 + \Delta]$ . This prior information can be approximated in the form of a model-based subspace,  $\langle \mathbf{W} \rangle \in \mathbb{G}(d_w, n)$ , which contains most of the energy of all steering vectors with DoAs in  $\Omega$ . The solution for the model-based subspace is given by the first  $d_w$  discrete prolate spheroidal wave functions [18], and the corresponding  $d_w$ -dimensional Slepian subspace [19], [20]. With this prior information, the problem amounts to estimating a subspace  $\langle \mathbf{V}_s \rangle \in \mathbb{G}(K, n)$  of dimension  $d_s = K$  that intersects  $\langle \mathbf{W} \rangle$  in  $d_i = 1$  dimension. As the average of projection matrices we use  $\bar{\mathbf{P}} = \hat{\mathbf{R}} / \text{tr}(\hat{\mathbf{R}})$ , where  $\hat{\mathbf{R}}$  is the sample covariance matrix estimated from  $M$  snapshots.

To evaluate the performance of the subspace estimation algorithms, we use the relative efficiency defined as

$$\eta = 1 - \|\mathbf{P}_V^\perp \mathbf{V}_0\|_F^2, \quad (13)$$

where  $\mathbf{V}_0$  is a unitary basis for the subspace spanned by the  $K$  true steering vectors, and  $\mathbf{P}_V^\perp$  is the projection onto the orthogonal complement of the estimated subspace. Note that

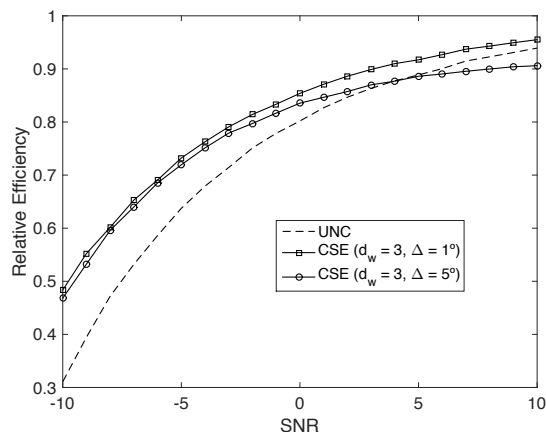


Fig. 3. Relative efficiency versus the SNR for the CSE algorithm and the unconstrained solution ( $n = 20$ ,  $d_w = 3$ ,  $d_s = 2$ ,  $d_i = 1$ ).

when  $\eta = 1$  the estimated subspace is perfectly aligned with the array manifold vectors.

We consider a scenario with  $n = 20$  antenna elements,  $K = 2$  sources with equal power, and  $M = 50$  snapshots. We assume that an estimate of  $\theta_1$  with a mismatch error  $\pm\Delta$  is available, and from this prior information we obtain a model-based subspace of dimension  $d_w = 3$ . The CSE algorithm estimates a subspace of dimension  $d_s = 2$  that intersects  $\langle \mathbf{W} \rangle$  in  $d_i = 1$  dimension. For comparison, we extract the 2-dimensional signal subspace of the sample covariance matrix (labeled as UNC in the figure). Fig. 3 shows the relative efficiency versus the SNR for two different values of the mismatch error. For low SNR values, the additional information provided by the model-based subspace provides a significant boost in performance even for very large values of the mismatch error. However, for high SNRs, the constrained solution is only better for small values of the mismatch error, since otherwise the prior information provided by the model-based subspace is not reliable and forcing an intersection can even be detrimental.

## V. CONCLUSIONS

Finding an average subspace that intersects another subspace in a given number of dimensions is an interesting optimization problem that arises in a number of applications. In this paper, we have used semidefinite relaxation techniques to form a convex relaxation of this problem, and then proposed the Constrained Subspace Estimation (CSE) algorithm. Some preliminary results in a beamforming problem indicate that, when the model-based subspace provides reliable prior information about the solution, the CSE algorithm outperforms the unconstrained estimate. Future work will include the analysis of the tightness of the relaxation, and a more detailed performance evaluation of the CSE algorithm in several application domains.

## ACKNOWLEDGMENT

This work has been supported by the Ministerio de Economía, Industria y Competitividad (MINECO) of Spain, and AEI/FEDER funds of the E.U., under grants TEC2013-47141-C4-R (RACHEL), TEC2016-75067-C4-4-R (CARMEN) and TEC2016-81900-REDT (KERMES), and by the National Science Foundation under Grant No. IIS-1633830.

## REFERENCES

- [1] P. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa, "Statistical computations on Grassmann and Stiefel manifolds for image and video-based recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2273–2286, Nov. 2011.
- [2] C. J. C. Burges, "Dimension reduction: a guided tour," *Foundations and Trends in Machine Learning*, vol. 2, no. 4, pp. 275–365, 2009.
- [3] R. R. Hagege and J. M. Francos, "Universal manifold embedding for geometrically deformed functions," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3676–3684, Jun. 2016.
- [4] I. Santamaria, L. L. Scharf, C. Peterson, M. Kirby, and J. Francos, "An order fitting rule for optimal subspace averaging," in *IEEE Statistical Signal Processing Workshop (SSP)*, 2016.
- [5] T. Marrinan, J. R. Beveridge, B. Draper, M. Kirby, and C. Peterson, "Finding the subspace mean or median to fit your need," in *Proc. of Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [6] A. N. Schwickerath, *Linear models, signal detection, and the Grassmannian manifold*. Ph.D. Thesis, Colorado State University, CO, USA, 2014.
- [7] T. P. Marrinan, *Grassmann, Flag, and Schubert Varieties in Applications*. Ph.D. Thesis, Colorado State University, CO, USA, 2016.
- [8] Z. Yavo, J. Francos, I. Santamaria, and L. L. Scharf, "Estimating the mean manifold of a deformable object from noisy observations," in *IEEE Image Video and Multidimensional Signal Processing Workshop (IVMSP)*, 2016.
- [9] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Proc. Magazine*, vol. 13, no. 4, pp. 67–94, Aug. 1996.
- [10] A. Srivastava and E. Klassen, "Monte Carlo extrinsic estimators of manifold-valued parameters," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 299–308, Aug. 2002.
- [11] A. Edelman, T. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
- [12] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.
- [13] S. Shahbazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal spacetime block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506–4517, Dec. 2005.
- [14] F. Hlawatsch and W. Kozek, "Time-frequency projection filters and time-frequency signal expansions," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3321–3334, Dec. 1994.
- [15] Z. Q. Luo, W. K. Ma, A. M. So, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Proc. Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.
- [16] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.0 beta*. <http://cvxr.com/cvx>, 2013.
- [17] Y. Huang and D. P. Palomar, "Randomized algorithms for optimal solutions of double-sided QCQP with applications in signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1093–1108, Mar. 2014.
- [18] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty-v: The discrete case," *Bell Syst. Tech. J.*, pp. 1371–1430, 1978.
- [19] O. Besson, L. L. Scharf, and F. Vincent, "Matched direction detectors and estimators for array processing with subspace steering vector uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4453–4463, Dec. 2005.
- [20] A. Pezeshki, B. D. Van Veen, L. L. Scharf, H. Cox, and M. L. Norderivaad, "Eigenvalue beamforming using multirank MVDR beamformer and subspace selection," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1954–1966, May 2008.